

### **REMARKS**

Claims 1-8 are pending. Claims 1, 3-5, and 7 has been amended to more distinctly describe the invention.

#### Rejection under 35 U.S.C. §112

On page 2, paragraph 1 of the office action dated May 17, 2007, the Examiner has queried the definition of "intein" as presented on page 20 of the specification. The Examiner reports that the concept of self splicing of a native intein in a native context is confusing to the Examiner in view of the claimed thiol-induced cleavage of an intein defined on page 20 from an intein-fusion protein.

The carrier conjugates described here are orientated so that the C-terminal end of the carrier protein is fused to the N-terminal end of the intein (see Figure 1 in the Application). The intein characteristically has an N-terminal cysteine.

The completion of a splicing reaction by a native intein can be inhibited by a thiol reagent or by a mutation within the intein so that cleavage but not splicing occurs. In the presence of a thiol reagent, a C-terminal thioester is added to the C-terminal end of the carrier protein during cleavage.

A comparison of splicing and intein-mediated peptide ligation is provided in Figure 1 of the accompanying paper authored by Evans et al. (*Protein Science* 7:2256-2264 (1998) - attached hereto). It is well understood in the art that a thiol reagent does not effectuate self-splicing but rather inhibits splicing and promotes cleavage.

The Examiner recited the first line of a definition that extends for 10 lines. In fact, the definition further states:

An intein further includes derivatives or modifications of the listed inteins including amino acid substitutions as long as these derivatives and modifications permit cleavage...It is known that certain mutations do not prevent an intein-induced cleavage reaction. (Application, page 20)

The specification then describes how:

a carrier protein or a ligand protein is positioned upstream of the intein (for example, a vector selected from Impact™ or designed according to an IMPACT™ vector design. (Application, page 30, line 27)

In paragraph 2 of the office action dated May 17, 2007, the Examiner has confused the N-terminal cysteine of the intein with the C-terminal thioester on the carrier protein. The formation of a C-terminal thioester on the carrier protein is the result of cleavage of an intein fusion protein in the presence of a thiol reagent. In the presence of a thiol reagent (e.g. 2-mercaptoethanesulfonic acid: MESNA), the intein is cleaved at its N-terminal end as shown in Figure 1 in the above application and a thioester is added to the C-terminal end of the carrier protein. The N-terminal cysteine referred to by the Examiner in paragraph 2 of the office action date May 17, 2007 is the first amino acid on the intein, **not** the protein.

Intein-mediated cleavage to generate a C-terminal thiol reagent on a protein in the presence of MESNA is the subject of U.S. patent 6,849,428 assigned to New England Biolabs, Inc. (NEB). When describing intein-mediated ligation in for example column 13 of the reference, Knock et al. refer to IMPACT™ (NEB) (see attached product

information from the NEB catalog dated 2002/2003), and to U.S. 5,834,247 (also assigned to NEB and authored by the present inventors).

The Examiner has additionally rejected the claims as indefinite in view of the use of the term "intein" and "carrier-intein fusion protein" and the role of MESNA because of the apparent confusion by the Examiner with respect to self splicing and cleavage. The discussion in the above paragraphs addresses this point. In view of the above, Examiner is respectfully requested to reverse the rejection.

The Examiner refers to creating a reactive C-terminal thioester on one of the carrier or ligand as a product of intein cleavage in the presence of a thiol reagent as mechanistically false. Applicants respectfully submit that as explained above, a protein fused to an intein according to the description in the application may be cleaved from the intein in the presence of MESNA to produce a reactive C-terminal thioester. In the present claimed invention, the carrier protein (not the ligand) is cleaved from the intein to generate a C-terminal thioester.

The description for Figure 1 in the Application states:

to induce cleavage of the protein bond prior to the N-terminal cysteine of the intein resulting in the formation of a reactive thioester on the C-terminus of the wild type CBD

This embodiment is consistent with the explanation provided above and is an example of the claimed invention where claim 1 as amended requires that a "a carrier-intein fusion protein is cleaved in

the presence of 2-mercaptoethanosulfonic acid to generate the C-terminal thioester on the carrier protein”.

Rejection under 35 U.S.C. §102

The Examiner has rejected the claims as anticipated by Knock et al. The Knock reference does not anticipate each and every element of the claimed invention and therefore does not anticipate the claimed method under 35 U.S.C. §102. The Knock reference describes how to immobilize proteins using intein-mediated ligation between a peptide and an anchor protein. In contrast, the present application utilizes intein-mediated ligation as a step in the purification of a wide range of ligand-binding molecules which have the property of binding to an immobilized ligand.

<b>Present application</b>	<b>Knock et al.</b>
Carrier: C-terminal thioester on the Carrier (anchor) must be a protein. Pages 14-15	Anchor: Nucleophilic group on the Anchor (carrier) – anchor does not need to be a protein. ¶ 0023
Ligand: Nucleophilic group on the Ligand. Ligand does not need to be a protein. Pages 14-15	Ligand: C-terminal thioester on the ligand must be a protein. ¶ 0023
Carrier binds to matrix	Anchor binds matrix
Ligand reacts with ligand-binding molecule.	Peptide reacts with peptide-binding molecule.
Elute ligand-binding molecule from ligand under conditions where the ligand-carrier remains bound to the matrix.	No elution step described. Only immobilization.

There are significant differences between Knock et al. and the present application.

The present claimed method is an improvement to Knock et al.

Knock et al. require that the ligand be a protein/polypeptide. This is explicitly stated in for example: paragraphs 0012-0019, 0021, 0042, 0045, 0047, 0048, 0049, 0066 and claims. Even if it were not explicitly stated, the fact that the protein/polypeptide carries the C-terminus thioester requires that it be synthesized as a fusion protein where the reactive thioester results from intein cleavage.

The Knock reference is additionally a method for immobilizing proteins and not a method for purifying protein that requires an elution step. There is no suggestion or requirement in Knock et al. that a molecule binding to the polypeptide be capable of elution or more particularly capable of elution without causing the anchor or bound peptide to become disassociated from the substrate.

In contrast, the present claimed method does not require that the ligand be a protein because it does not have the reactive thioester at its C-terminus. Instead it has an N-terminus that is a nucleophile. Consequently, the ligand need not be expressed as a fusion protein. In fact the specification states that the ligand **may** be a protein, DNA, a carbohydrate or a small molecule. (Application, page 19 line 5).

The present claimed method is additionally a method for purifying a ligand-binding molecule from a mixture. This method requires an elution step. The method also requires that the carrier be associated with a matrix non-covalently.

Consequently, the Knock et al. reference does not anticipate the claimed invention.

### **CONCLUSION**

For the reasons set forth above, Applicants respectfully request that the rejections set forth in the Final Office Action of May 17, 2007 be withdrawn and submit that this case is in condition for immediate allowance. Early and favorable consideration leading to prompt issuance of this Application is earnestly solicited.

Applicants are concurrently filing a notice of appeal and a petition for a three-month extension of time to file a response. Applicants authorize that the fees for the notice and extension, which total \$780, be charged to Deposit Account No. 14-0740. Applicants also authorize that any deficiencies be charged to Deposit Account No. 14-0740.

Respectfully submitted,

NEW ENGLAND BIOLABS, INC.

Date: November 14, 2007

Customer No.: 28986

/Harriet M. Strimpel, D.Phil/  
Harriet M. Strimpel, D.Phil.  
(Reg. No.: 37,008)  
Attorney for Applicant  
240 Country Road  
Ipswich, Massachusetts 01938  
(978) 380-7373

EVANS ET AL. PROTEIN SCIENCE 7: 2256-2264 (1998)



# Protein Science

## Semisynthesis of cytotoxic proteins using a modified protein splicing element

T. C. EVANS-JR., J. BENNER and M. Q. XU

*Protein Sci.* 1998 7: 2256-2264

---

### References

Article cited in:  
<http://www.proteinscience.org/cgi/content/abstract/7/11/2256#otherarticles>

### Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#)

---

### Notes

---

To subscribe to *Protein Science* go to:  
<http://www.proteinscience.org/subscriptions/>

---

## Semisynthesis of cytotoxic proteins using a modified protein splicing element

THOMAS C. EVANS, JR., JACK BENNER, AND MING-QUN XU

New England Biolabs Inc., 32 Tozer Road, Beverly, Massachusetts 01915

(RECEIVED April 9, 1998; ACCEPTED June 15, 1998)

### Abstract

Two cytotoxic proteins, bovine pancreatic ribonuclease A (RNase A), and a restriction endonuclease from *Haemophilus parainfluenzae* (*Hpa*I), were produced using a novel semisynthetic approach that utilizes a protein splicing element, an intein, to generate a reactive thioester at the C-terminus of a recombinant protein. Nucleophilic attack on this thioester by the N-terminal cysteine of a synthetic peptide ultimately leads to the ligation of the two reactants through a native peptide bond. This strategy was used to produce RNase A and *Hpa*I by isolating inactive truncated forms of these proteins, the first 109 and 223 amino acids of RNase A and *Hpa*I, respectively, as fusion proteins consisting of the target protein, an intein, and a chain binding domain. Thiol-induced cleavage of the precursor led to the liberation of the target protein with a C-terminal thioester tag. Addition of synthetic peptides representing the amino acids missing from the truncated forms led to the generation of full-length products that displayed catalytic activity indicative of the wild-type enzymes. The turnover numbers and  $K_m$  for ligated and renatured RNase A were  $8.2 \text{ s}^{-1}$  and 1.5 mM, in good agreement with reported values of  $8.3 \text{ s}^{-1}$  and 1.2 mM (Hodges & Merrifield, 1975). Ligated *Hpa*I had a specific activity of  $0.5\text{--}1.5 \times 10^6 \text{ U/mg}$ , which compared favorably with the expected value of  $1\text{--}2 \times 10^6 \text{ U/mg}$  (J. Benner, unpubl. obs.). Besides assisting in the production of cytotoxic proteins, this technique could allow the easy insertion of unnatural amino acids into a protein sequence.

**Keywords:** cytotoxic proteins; intein; protein ligation; protein splicing element; semisynthesis

The engineering and study of proteins have greatly benefited from techniques that allow the manipulation of amino acid composition at the protein level. These methodologies broaden the realm of protein engineering to include the incorporation of unnatural amino acids and the inclusion of proteins that are cytotoxic or difficult to express (Offord, 1987; Kent, 1988; Roy & Acharya, 1994; Wallace, 1995). These include *in vitro* translation using chemically misacylated tRNAs, total chemical synthesis of polypeptides, and protein semisynthesis. The use of an *in vitro* translation system bypasses problems associated with toxicity and restriction to coded amino acids, but the final protein yields tend to be low (Noren et al., 1989). Solid-phase total chemical synthesis involves the stepwise addition of amino acids to a peptide anchored to a support

through its C-terminal residue (Kent, 1988). Unlike proteins expressed *in vivo*, synthesized peptides are not restricted to the coded amino acids. Unfortunately, a major limitation of this technology is that polypeptides of larger than about 100 amino acids are difficult to obtain.

To overcome this limitation and to develop the potential for more extensive manipulation of proteins, new methods that lead to the precise condensation of synthetic peptides (Dawson et al., 1994; Jackson et al., 1994; Liu & Tam, 1994) and protein semisynthesis (Offord, 1987; Roy & Acharya, 1994; Wallace, 1995) have become the focus of many studies. Semisynthesis can be achieved by several approaches, and results in the fusion of a chemically manipulated or synthetic peptide to other fragments of a protein. Studies on cytochrome *c* (Wallace, 1993) and insulin (Nakagawa & Tager, 1986) utilizing semisynthesis to incorporate noncoded amino acids have provided insights into their structure/function relationships. However, current semisynthetic techniques can be limited by protein size, the requirement for proteases, and/or the need for specialized reaction conditions.

The ligation methodology presented in this paper takes advantage of an intein whose protein splicing activity has been blocked by mutation. Inteins are naturally occurring proteins that catalyze their own excision from a precursor protein with the simultaneous ligation of two flanking protein sequences, resulting in the generation of two

Reprint requests to: Ming-Qun Xu, New England Biolabs Inc., 32 Tozer Road, Beverly, Massachusetts 01915; e-mail: xum@neb.com.

**Abbreviations:** CBD, chain binding domain; CTP, cytidine 2',5'-cyclic monophosphate; DTT, dithiothreitol; GSH, reduced glutathione; GSOG, oxidized glutathione; *Hpa*I, a restriction endonuclease from *Haemophilus parainfluenzae*; *Hpa*I<sub>109</sub>, the first 109 amino acids of *Hpa*I; MBP, maltose binding protein; MESNA, the sodium salt of 2-mercaptoethanesulfonic acid; MYB, a fusion protein of MBP-*See* VMA intein (N454A)-CBD; MKB, a fusion protein of MBP-*Mae* GyrA intein (N198A)-CBD; RNase A, bovine pancreatic ribonuclease A; RNase A<sub>109</sub>, the first 109 amino acids of RNase A with a methionine added to the N-terminus.

proteins from a single gene product (Fig. 1A; Kane et al., 1990; Perler et al., 1994). The splicing reaction is initiated at the N-terminal residue of the intein with an N-S acyl-shift (in inteins with an N-terminal cysteine) and release of the intein from the precursor is mediated by the cyclization of a C-terminal asparagine residue (Xu et al., 1994; Shao et al., 1995; Chong et al., 1996; Xu & Perler, 1996).

Mutation of the reactive asparagine residue to alanine blocks splicing by preventing cleavage at this splice junction. However, this mutant can still undergo the N-S shift, making it sensitive to thiol-mediated cleavage between itself and an N-terminal target protein (Chong et al., 1997). In practice, the target protein is isolated from *Escherichia coli* cell extract as a three-part fusion protein consisting of the target protein, an engineered intein, and a chitin-binding domain (CBD) affinity tag (Fig. 1B). Thiol-induced cleavage of the fusion protein results in the release of the target protein with a C-terminal thioester (Chong et al., 1997). A previously described protocol, known as "native chemical ligation" (Dawson et al., 1994), resulted in the fusion of two synthetic peptides when the N-terminal cysteine of one peptide attacks a C-terminal thioester present on another peptide. Initially, a new thioester bond is formed between the peptides, but this spontaneously rearranges to generate a stable peptide bond. Interestingly, this process is analogous to steps 2–4 of the protein-splicing pathway (Fig. 1A). The potential of a procedure such as "native chemical ligation" can be greatly increased by taking advantage of the thioester produced on a bacterially expressed protein during intein-mediated purification.

In this paper, we describe a new method for protein semisynthesis that allows the facile production of a recombinant protein fragment fused to a synthetic peptide through a native peptide bond. This technique was used to produce two potentially cytotoxic proteins, bovine pancreatic ribonuclease A (RNase A) and the type II restriction endonuclease *HpaI* (Ito et al., 1992), by first expressing inactive fragments in bacteria followed by reaction with synthetic peptides to generate enzymatically active fusion products.

## Results

### Generation of thioester-tagged proteins

The ligation reaction involves the complementing of two reactive groups, the C-terminal thioester of a protein and the N-terminal cysteine of a synthetic peptide (Fig. 1B). Proteins with C-terminal thioesters were generated as enzymatically inactive, truncated forms of bovine pancreatic ribonuclease A and *HpaI* using the *Mxe* GyrA intein (Telenti et al., 1997), with an asparagine 198 to alanine mutation (N198A). This protocol was similar to a purification technique described previously for a modified *See* VMA intein (Chong et al., 1997). The truncated proteins, the first 109 amino acids of RNase A (RNase A<sub>109</sub>) with a methionine added to the N-terminus to aid in bacterial expression and the first 223 amino acids of *HpaI* (HpaI<sub>223</sub>) were expressed in *E. coli* as three-part fusion proteins consisting of the target protein, the intein, and the CBD (Figs. 1B, 3, 4). The fusion protein was isolated by binding to a column containing chitin beads. Thiol-induced cleavage released the target protein with a thioester at the C-terminal alpha carbon.

Previously, the nature of the C-terminal thioester was found to influence the rate of ligation for "native chemical ligation" (Dawson et al., 1994). In this study, the thiol used to induce intein-mediated cleavage had a profound effect on ligation efficiency. Tests were con-

ducted with amino acids known to be good for thiol-induced cleavage, tyrosine, and glycine for the *Mxe* GyrA (N198A) and *See* VMA (N454A) inteins, respectively, at the C-terminus of maltose binding protein (MBP, New England Biolabs, Beverly, Massachusetts) (Duplay et al., 1984; Guan et al., 1987) from *E. coli* (Fig. 2A). Two of the most effective reagents at inducing intein-mediated cleavage, dithiothreitol and 2-mercaptoethanol (Chong et al., 1997), displayed low ligation efficiency (20–30%, see Fig. 2A,B). Other compounds were tested for both the ability to induce intein-mediated cleavage and promote efficient ligation. Two compounds were found to display >80% ligation efficiency after an overnight incubation at 4 °C, the sodium salt of 2-mercaptoethanesulfonic acid (MESNA, Sigma Chemical Co., St. Louis, Missouri) and thiophenol (Aldrich Chemical Co., Milwaukee, Wisconsin) (Fig. 2B). Thiophenol had been used in the past with "native chemical ligation" to increase the rate of ligation (Lu et al., 1996). However, neither compound caused more than 30% cleavage of the *See* VMA intein fusion protein after overnight incubation at 4 °C (Fig. 2A). In contrast, the modified *Mxe* GyrA intein displayed efficient cleavage (>90%) after overnight incubation at 4 °C when either of these compounds were present at a concentration of 50 mM (Fig. 2A). MESNA is odorless, and was used in future experiments.

In addition to the sensitivity of different inteins to different thiols, thiol-induced cleavage of the same intein can be dramatically altered by the target protein C-terminal residue both in the *See* VMA intein (Chong et al., 1998) and the *Mxe* GyrA intein (F.B. Perler and M.W. Southworth, pers. obs.). In this paper, the *Mxe* GyrA (N198A) intein did not cleave efficiently with the alanine (Ala<sub>109</sub>) of the truncated RNase A adjacent to the first residue of the intein, and led to a lower than expected yield of RNase A (about 0.3 mg/L cell culture after HPLC purification, whereas, based on estimates of precursor amounts, the expected yield would be 4–8 mg/L cell culture). However, the *HpaI* fragment had a methionine (Met<sub>223</sub>) adjacent to the first residue of the intein and it cleaved efficiently. This resulted in the isolation of 3–6 mg/L cell culture of ligated *HpaI*. In fact, in cases where maximum yield is desired it may be advantageous to carefully select the intein, thiol, and target protein.

### Protein ligation

The thioester-tagged proteins were allowed to undergo ligation with a peptide containing an N-terminal cysteine by adding lyophilized peptide (1 mM final concentration) to the target protein fraction and incubating overnight at 4 °C. Ligation efficiency was high for both truncated proteins (>90%, Fig. 4) when MESNA was used. Furthermore, the condensation reactions occurred under a variety of conditions. Intein-mediated peptide ligation, with MBP, has been observed in Tris buffers (50 mM Tris-HCl, 100 mM NaCl) in which the pH varied from 6.0 to 9.0, although the efficiency of ligation was lower at pH 6 (>60%) than at pH 9.0 (>90%). Ligation also proceeds well (>80%) in Tris buffers (50 mM Tris-HCl, pH 8.0) containing 0.1–1 M NaCl or 6 M urea (data not shown). Also, the small peptide can be synthesized with unnatural amino acids, radioactive residues, or other properties not possible from recombinant sources. For example, using this procedure a biotinylated peptide has been ligated to MBP (data not shown). A limitation, as with all semisynthetic methods, is that the residue to be modified must occur within a peptide of a size that can be chemically synthesized.

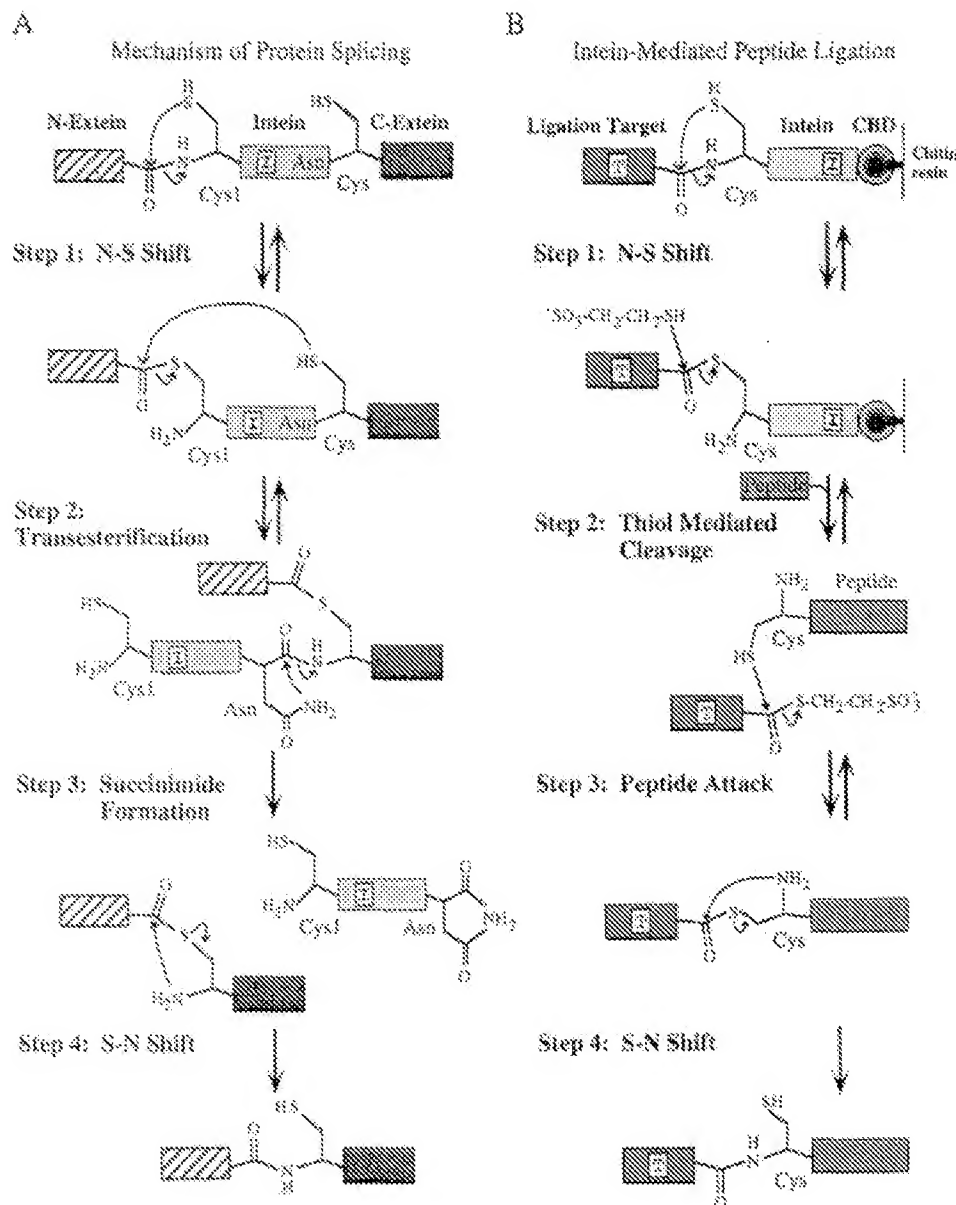


Fig. 1. Mechanism of protein splicing and peptide ligation. **A:** Protein splicing of an intein with an N-terminal cysteine. Step 1 is an N-S acyl shift that occurs between the C-terminal amino acid of the N-extein (hatched box) and the N-terminal cysteine of the intein (yellow box) and results in the formation of a reactive thioester. This thioester is the focus of a transesterification (step 2) reaction that results in a branched intermediate at the N-terminal cysteine of the C-extein (hatched box). The branched intermediate is resolved by the cyclization of an asparagine residue at the C-terminus of the intein to form a succinimide group (step 3). The intein is released from the precursor and the extein residues are attached through a thioester bond. A spontaneous S-N acyl shift generates a peptide bond between the extein sequences and results in a fusion product with a peptide bond at the site of condensation (step 4). **B:** Intein-mediated peptide ligation. As in protein splicing, the initial step is an N-S acyl shift that generates a thioester between the target protein (red box) and the intein (yellow box, step 1). The chitin binding domain from *B. subtilis* (orange box) [Saito, Watanabe et al., 1994] allows easy purification of the three-part fusion protein by interaction with chitin resin (blue circle). Unbound proteins are washed away. A nucleophilic attack on the thioester by a small thiol compound (step 2) cleaves the precursor protein and generates a new thioester at the C-terminus of the target protein. The cleaved target protein is eluted from the chitin resin, while the intein-CBD remains bound. The newly formed thioester is reactive to attack by a cysteine residue at the N-terminus of a peptide (green box, step 3). In common with step 3 of protein splicing, this results in a thioester bond between the target protein and the peptide. An S-N acyl shift results in the formation of a peptide bond between the peptide and the target protein. It should be noted that although the direct attack of a peptide on the thioester formed between the target protein and the intein is theoretically possible, it was not found to occur under these conditions (data not shown).

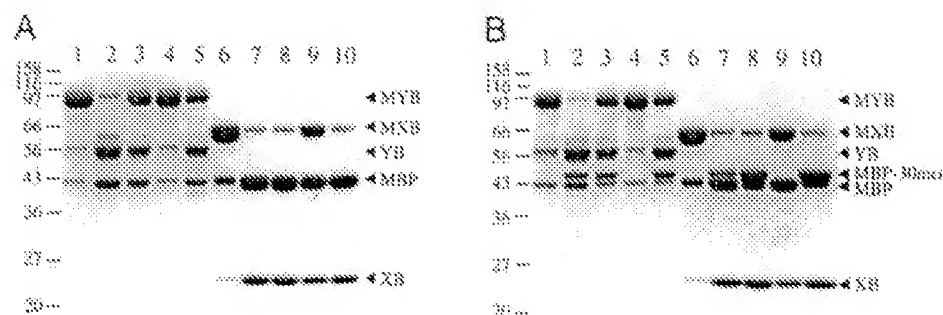


Fig. 2. Cleavage and ligation reactions with different thiols visualized on 10–20% w/v gels. A: Thiol-induced cleavage of MYB and MXB. MYB (a fusion protein of maltose binding protein–*See VMA* intein (N434A)–chitin binding domain) and MXB (a fusion protein of maltose binding protein–*Myx* (N100A)–chitin binding domain) were incubated overnight at 4°C with various thiols (50 mM) in 150 mM Tris, 100 mM NaCl, pH 8 as described in Materials and methods. Lanes 1–5: cleavage of MYB. Lane 1, no thiol; lane 2, dithiothreitol; lane 3, 2-mercaptoethanesulfonic acid; lane 4, 3-mercaptopyropionic acid; lane 5, thiophenol. Lanes 6–10: cleavage of MXB. Lane 6, no thiol; lane 7, dithiothreitol; lane 8, 2-mercaptoethanesulfonic acid; lane 9, 3-mercaptopyropionic acid; lane 10, thiophenol. B: Thiol-mediated cleavage and ligation with MYB and MXB. MYB and MXB were incubated as described in A except that a 30-amino acid peptide was also present (1 mM) (see Materials and methods). The peptide ligates to the C-terminus of maltose binding protein. Lanes 1–5: ligation with MYB. Lane 1, no thiol; lane 2, dithiothreitol; lane 3, 2-mercaptoethanesulfonic acid; lane 4, 3-mercaptopyropionic acid; lane 5, thiophenol. Lanes 6–10: ligation with MXB. Lane 6, no thiol; lane 7, dithiothreitol; lane 8, 2-mercaptoethanesulfonic acid; lane 9, 3-mercaptopyropionic acid; lane 10, thiophenol.

The choice of a site to split a protein is constrained mostly by the presence of a cysteine, although mutating a residue to a cysteine to facilitate ligation may not alter the protein's activity. The residue C-terminal to the fusion site may affect ligation as discussed by Dawson et al. for native chemical ligation (Dawson et al., 1994). However, this study indicated that different amino acids at the C-terminus of the recombinant protein do not seem to significantly change the ligation efficiency, although this is based on a limited sampling. Fragments of RNase A and *HpaI* demonstrate efficient fusion with alanine and methionine at their C-termini. Furthermore, studies with MBP indicate that ligation also proceeds

well with leucine, glycine, or tyrosine as the C-terminal residue (data not shown).

#### Enzymatic activity of ligated proteins

The truncated fragments of RNase A and *HpaI* lacked any detectable enzymatic activity (data not shown for truncated RNase A); however, upon ligation with the appropriate synthetic peptides, both displayed catalytic activities typical of the wild-type enzymes (Fig. 5). Ligated RNase A required renaturation, as described previously (Tyler & Gilbert, 1991), to display full activity as determined by the catalyzed hydrolysis of 2',3'-cyclic cytosine monophosphate (Hodges & Merrifield, 1975). The ligated and renatured protein had  $k_{cat}$  and  $K_m$  values of  $8.2 \text{ s}^{-1}$  and 1.5 mM, respectively, in good agreement with previously reported values of  $8.3 \text{ s}^{-1}$  and 1.2 mM (Hodges & Merrifield, 1975). Final yields of truncated, ligated, and enzymatically active RNase A were about 0.3 mg/L cell culture for each, demonstrating the high efficiency of ligation.

In contrast to RNase A, *HpaI* displayed activity immediately after ligation. Specific activity of the recovered soluble enzyme was determined to be  $0.5\text{--}1.5 \times 10^6 \text{ U/mg}$ , in good agreement with the expected value of  $1\text{--}2 \times 10^6 \text{ U/mg}$  (J. Benner, unpubl. obs.). However, a significant portion of the ligated protein became insoluble and precipitated from solution, accounting for 95–98% of the ligation product as shown by SDS-PAGE (see Fig. 4). It appears, therefore, that only a fraction of the fusion product folded correctly while the rest assumed a misfolded conformation.

Unlike with RNase A, refolding experiments with *HpaI* were unsuccessful. A need for refolding has been found with other condensation techniques (Jackson et al., 1994; Lu et al., 1996) and is a potential complication of any method that involves the formation of a protein from smaller fragments. The final yields of truncated, ligated, and enzymatically active *HpaI* were 3–6, 3–6, and 0.1–0.2 mg/L cell culture, respectively. Incubation of the thioester-tagged protein fragments of RNase A and *HpaI* with hydroxylamine, which replaces the thioester with a hydroxamate that is unreactive to

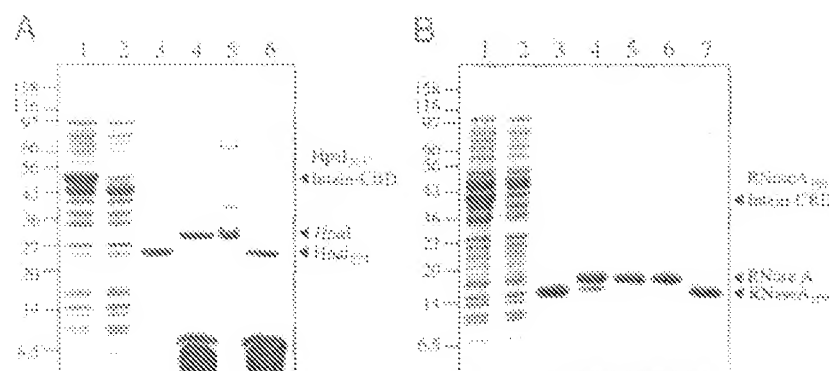
#### Bovine pancreatic RNase A

1 MKETAAARKE RQHMDSTSA ASLANVYCHQM AKSRMLTKDR  
41 CKPQNTFVIR SLAIVQAVS QKQVACKRQY YNCYCRYSYIN  
81 STUCKREIS SRYPNCAVET FGANKHVA CEUSXPYTHHDAK

#### *HpaI*

1 MKYTERSRV FVUSPYFNY SQCVLRIVE IUNGKIMED DRHIVNLS  
51 RGLPLESINK IAFIMKAWA EVFQSDIN RQRYNLSV EAOEELDSI  
101 MDAQYNNEN YTCNVYDA LSNDPDSOK SPNTHSR EYAYVKIDPF  
151 WHSIRKS VYVKRNYTN LADGIMQHD FVYERKYIS LDEIVNPAI  
201 QUGQDQKH HYVYSQRTY WQRCALILE YLRKRTIE QYNEAKRKE  
251 SHKQ

Fig. 3. Amino acid sequences of bovine pancreatic RNase A and *HpaI*. Note that a methionine has been added to the N-terminus of RNase A to allow translation in the bacterial system. RNase A and *HpaI* were terminated at residue 109 (RNase A<sub>109</sub>, native RNase A numbering) and methionine 223 (*HpaI*<sub>223</sub>), respectively, because both of these residues were followed by naturally occurring cysteines. The remaining peptides (underlined) were synthesized by standard solid-phase procedures (New England Biolabs).

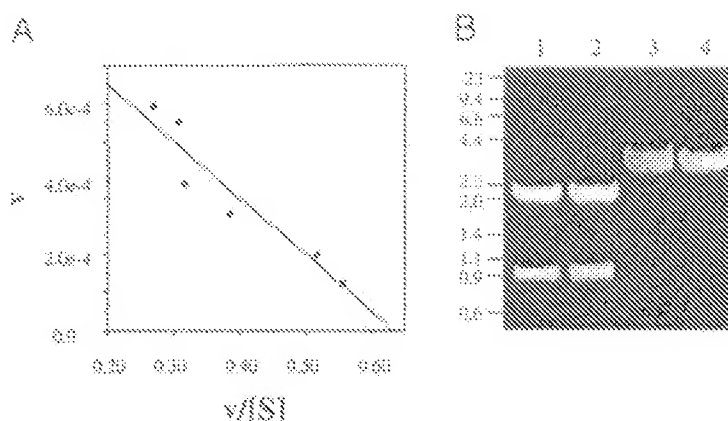


**Fig. 4.** Peptide ligation reactions examined using SDS-PAGE on 10–20% tricine gels. **A:** Lane 1: clarified cell extract after IPTG (0.5 mM) induction of EB2566 cells containing the pTXB1-Hpa1 plasmid. The fusion protein of Hpa1<sub>123</sub>-Myc GyrA Intein-CBD (55 kDa) is visible. Lane 2: cell extract as in lane 1 after passage over a chitin column, which results in the binding of the fusion protein. Lane 3: Hpa1<sub>123</sub> (25.7 kDa) after cleavage from the fusion protein by addition of MDESA. Lane 4: ligation product of Hpa1<sub>123</sub> (0.2 mg/mL) with 1 mM of a 31 amino acid peptide (ligation product 29.6 kDa), representing the residues necessary to generate full length Hpa1, after overnight incubation at 4 °C. Lane 5: full-length Hpa1 from a recombinant source (79.6 kDa) containing BSA (66 kDa) and two ligandless. Lane 6: hydroxylamine treated Hpa1<sub>123</sub> (overnight incubation at 4 °C with 30 mM hydroxylamine) reacted overnight at 4 °C with the 31 amino acid peptide. The low molecular weight smears present in lanes 4 and 6 are the 31 amino acid peptide used in ligation. **B:** Lane 1: clarified cell extract after IPTG induction of EB2566 cells containing the pTXB1-RNaseA plasmid. Lane 2: cell extract as in lane 1 after passage over a chitin column, which results in the binding of the fusion protein. Lane 3: RNase A<sub>123</sub> (12 kDa) after purification and concentration. Lane 4: ligation product of RNase A<sub>123</sub> (12 kDa) after incubation of RNase A<sub>123</sub> with a 15-mer peptide representing the missing amino acids (1 mM final peptide concentration). Lane 5: ligated bovine pancreatic RNase A after HPLC purification. Lane 6: full-length bovine pancreatic RNase A from the Sigma Chemical Co. Lane 7: RNase A<sub>123</sub> pretreated with hydroxylamine (30 mM hydroxylamine for 16 h at 4 °C) and further incubated with the 15-mer peptide for another 16 h at 4 °C.

nucleophilic attack by thiol-containing compounds, eliminated ligation and subsequent catalytic activity (Figs. 4, 5). The activity of full length RNase A and Hpa1 as well as previously ligated samples, were not affected by incubation with hydroxylamine.

## Discussion

The application of molecular biological approaches in the expression and modulation of cytotoxic proteins can be difficult. It is



**Fig. 5.** Catalytic activities of ligated RNase A and Hpa1. **A:** Activity of ligated bovine pancreatic RNase A toward the hydrolysis of cytidine 2,3'-cyclic monophosphate (C>P). Hydrolysis of C>P by ligated RNase A (0.61  $\mu$ M) was monitored by a change in absorbance at 296 nm.  $v$  (velocity) has units of  $\Delta A296/s$  and  $[S]$  is in mol/L. Each point represents the average of three separate determinations. The turnover number and  $K_m$  were determined to be  $8.2 \times 10^4$  and 1.5 mM, respectively. **B:** Endonuclease activity of ligated Hpa1 toward LITMUS 28 DNA (New England Biolabs). The assay consisted of incubation of the appropriate sample with Hpa1 (1 U, New England Biolabs) and 2  $\mu$ g of LITMUS 28 DNA overnight at 37 °C in 20 mM Tris-acetate, pH 7.9, 10 mM magnesium acetate, 50 mM potassium acetate, 1 mM dithiothreitol, and 120  $\mu$ g/mL BSA (total volume 30  $\mu$ L). The reactions were visualized on a 1.2% agarose gel stained with ethidium bromide. Lane 1: LITMUS 28 DNA incubated with 0.3 U of Hpa1 ligation product. Lane 2: LITMUS 28 DNA and 2.5 U of Hpa1 (New England Biolabs). Lane 3: LITMUS 28 DNA incubated with the equivalent amount of Hpa1<sub>123</sub> and 31 amino acid peptide as in lane 1. Ligation of these materials was prevented by incubating Hpa1<sub>123</sub> with hydroxylamine at 4 °C overnight before another 4 °C overnight incubation with the peptide. Hydroxylamine was not found to effect the activity of full length Hpa1. Lane 4: Ligated LITMUS 28 DNA.



often necessary to develop unique conditions that allow the production of a particular cytotoxic protein. Unfortunately, these special conditions usually are not applicable to another protein of interest. The procedure described here may allow the production of cytotoxic proteins as either wild-type or mutant forms. Moreover, because it relies on the inactivation of a protein by truncation, it may be more generally applicable. Both full-length RNase A and *HpaI*, in the absence of its respective methylase, are potentially toxic to cells when expressed at high levels, whereas the target-intein-CBD fusion proteins of both RNase A and *HpaI* could be expressed at  $\geq 5$  mg/L of cell culture and converted to full length, enzymatically active proteins by fusing the thioester-tagged fragments to synthetic peptides representing the missing amino acids.

The nature of the C-terminal thioester influenced ligation efficiency as was seen previously (Dawson et al., 1994; Lu et al., 1996), and the thiol reagent used for cleavage determines the properties of the C-terminal thioester. Therefore, the protocol described in this paper has two important requirements for a thiol reagent. First, it must promote efficient ligation. Two compounds shown in Figure 2B have this property—MESNA and thiophenol. Second, total yields of thioester-tagged proteins would be higher if the thiol reagent proficiently induced intein-mediated cleavage. Neither MESNA nor thiophenol-induced efficient cleavage at 4°C with the *Sec* VMA intein fusion protein (Fig. 2A). Although cleavage was more efficient (30–60%) with MESNA at room temperature (data not shown). However, another intein, a modified *Mxe* GyrA intein, did cleave efficiently with both of these compounds (Fig. 2A) and was used in subsequent purification reactions. The differential sensitivity of inteins to thiol reagents may indicate structural variations that allow the active site in one intein to be more solvent accessible. However, both of the MBP fusions of the modified *Sec* VMA and *Mxe* GyrA inteins are resistant to hydrolysis (data not shown). Therefore, the ability of a thiol reagent to induce cleavage may be due to the correct spatial arrangement of atoms that allows the thiol group to orient itself in the optimum position for nucleophilic attack in one intein but not another.

Similar to chemoselective techniques, which rely on the pairing of two reactive groups, this strategy allows ligation with no need for the presence of protecting groups on either the protein fragment or the synthetic peptide. Previously, reactive groups have been introduced into the peptide during synthesis and/or into the protein fragment chemically or with proteases (Wallace, 1995). An advantage of the procedure presented in this paper is that a reactive thioester is easily and precisely generated on a heterologously expressed protein by the intein as part of the purification process. Furthermore, this bypasses problems associated with proteases such as the presence of multiple protease sites within a protein and the need for ligation to occur at a protease recognition sequence. However, there is a requirement for a cysteine at the site of ligation. Application of mutagenesis to insert ligation sites may be useful to increase the number of proteins accessible to this procedure, as was demonstrated for "autocatalytic fragment religation" involving cytochrome *c* (Woods et al., 1996).

The ligation of a synthetic peptide to the C-terminus of a recombinant protein complements other recent techniques such as native chemical ligation (Dawson et al., 1994) and the use of "subtiligase" (Jackson et al., 1994), which can be used to attach compounds to the N-terminus of a protein sequence. A slight modification of native chemical ligation was used to attach an EDTA derivative to the first residue of chimeric *c-Fos* and *c-Jun* bZip proteins (Erlanson et al., 1996). Use of subtiligase has allowed the

ligation of peptides to the N-terminus of methionyl-extended human growth hormone and atrial natriuretic peptide (Chang et al., 1994). Therefore, modification of a protein sequence can be accomplished at either or both ends using these techniques in concert with the methodology described in this paper.

An important consideration when using this methodology is at which point to truncate the target protein. In the case of RNase A, this was facilitated by the vast amount of biochemical data on this protein. Previous studies had demonstrated the importance of the C-terminal residues for RNase A activity (Lin, 1970; Ginte et al., 1972; Wlodawer et al., 1988). However, most cytotoxic proteins of interest have little or no available biochemical information. To address this issue, we applied this technique to the type II restriction endonuclease, *HpaI*, about which structure/function relationships are not well determined. Truncation of *HpaI* at residue 223 led to an inactive fragment. Inactivation could be due to the deletion of critical residues or the misfolding of the truncated protein. A portion of the protein displayed activity immediately after ligation. However, a significant percentage did not, probably due to misfolding. Renaturation is probably a requirement for any technique that involves the assembly of proteins from smaller fragments and was successful with RNase A but not with *HpaI*. The need for renaturation may be eliminated if the protein were divided between domains that fold independently.

The presented procedure may allow the production of cytotoxic proteins and the introduction of noncoded amino acids into large protein sequences. Also, the ability to precisely place a reactive thioester at the C-terminus of a large protein paves the way for specific labeling of protein sequences without the use of chemicals that may derivatize important catalytic residues. In the future, replacement of the synthetic peptide with a recombinant protein should allow the fusion of large unrelated proteins.

## Materials and methods

### Creation of vectors pTXB1 and pTXB2

Asparagine 198 of the *Mxe* GyrA intein was mutated to alanine by linker insertion into the *XmnI* and *PstI* sites of pRexMIP-TyrXim 3Pdel to create pMXP1. The linker was composed of *mxe*#3 (5'-GGTTCGTCAGCCACGCTACTGGCCCTCACCAGTTGATAGCTGCA-3') and *mxe*#4 (5'-GCTATCAACCGGTGAGGCCAGTAGCGTGGCTGACGAACC-3'). Into pMXP1 another linker (composed of *mxe*#1 5'-TCGAATCTAGACATATGCCCATGGTGCCGCGCCGCTCAGAGGCTCTTCTGCATCACGGGAGATGCA-3' and *mxe*#2 5'-CTAGTGCATCTCCCGTATGACAGGAAGAGCCCTCGAGGCGGCGCCGCCACCCATGGCCATATGTCTAGAT-3') was inserted into the *XhoI* and *SpeI* sites to introduce a multiple cloning site (*XbaI*-*NdeI*-*NcoI*-*NotI*-*XhoI*-*SapI*) before the *Mxe* GyrA intein (pMXP2). The *NotI* to *AgeI* digest of pMXP2 was ligated into the same sites in pTYB1 (New England Biolabs, Beverly, Massachusetts) and the *NcoI* to *AgeI* fragment of pMXP2 was cloned into pTYB3 (New England Biolabs) to create plasmids pTXB1 and pTXB2, respectively.

### Isolation of truncated genes for *HpaI* and RNase A

The genes for truncated *HpaI* (*HpaI*<sub>1-223</sub>, amino acids 1–223) were isolated by the polymerase chain reaction (PCR) (Saiki et al.,

[1985] from a cloned full length *HpaI* gene (New England Biolabs) using oligonucleotides Hpa#4 (5'-GTGGAAGCTCTTCAAGCACATTGCCAGGTTTCTCTT-3') and Hpa#5 (5'-GTGGAATCATGAAATACGAAGAAATTAATTTTAA-3'). The PCR reaction solution contained Vent DNA polymerase buffer (New England Biolabs), 0.39  $\mu$ M primer, 80  $\mu$ M dNTP, 20 U/mL Vent DNA polymerase (New England Biolabs) and 500 ng/mL of the plasmid containing the *HpaI* gene. Amplification was performed on a Perkin-Elmer (Norwalk, Connecticut) Cetus thermal cycler at 94 °C for 45 s, 60 °C for 30 s, and 72 °C for 50 s for 20 cycles. The PCR product was transferred to a pTXB2 vector by shuffling into the *NcoI* and *SapI* sites, yielding pTXB2-HpaI. The gene for truncated RNase A (RNase A<sub>109</sub>, amino acids 1–109 of full length RNase A with a methionine added to the N-terminus) was isolated by PCR from a Bovine Genomic Library from adult milk liver (CLON-TECH Laboratories Inc., Palo Alto, California) using oligos RNase#1 (5'-GTGGAACATATGAAGGAACTGCAGCAGCCA-3') and R#2 (5'-GTGGAAGCTCTTCAAGCAAGCCACAATGATGTGTTTATT-3'). PCR was as described for *HpaI* with the following exceptions. Lambda library (2  $\mu$ L) was heated to 94 °C for 1 min prior to use. The PCR reaction solution contained magnesium-free Vent DNA polymerase buffer (New England Biolabs) to which MgSO<sub>4</sub> was added to a final concentration of 1 mM and 25 U/mL Taq DNA polymerase (Fisher Scientific, Pittsburgh, Pennsylvania). The thermal cycler conditions were 94 °C for 30 s, 52 °C for 30 s, and 72 °C for 45 s for 25 cycles. The PCR product was cloned into the *NdeI* and *SapI* sites of pTXB1, resulting in pTXB1-RNase. Oligonucleotide synthesis and sequencing of the cloned genes were performed by the Organic Division and Sequencing Facility of New England Biolabs, respectively.

#### Purification of thioester-tagged proteins

Protein purification was as described using the *Xcc* VMA intein (Chang et al., 1997) with slight modification. ER2566 cells (W. Jack, D. Dila, J. Menin, E.A. Raleigh, in prep.) containing the appropriate pTXB vector were grown to an OD<sub>600</sub> of 0.5–0.6 at 37 °C, at which point they were induced with 0.5 mM IPTG overnight at 15 °C. Cells were harvested by centrifugation and lysed by sonication (performed on ice). The three part fusion protein was bound to chitin beads (10 mL bed volume) equilibrated in Buffer A (50 mM Tris, pH 7.4, and 500 mM NaCl), and washed with 10 column volumes of Buffer A to remove unbound material. Cleavage was initiated using a buffer of 50 mM MESNA (Sigma Chemical Co., St. Louis, Missouri), 50 mM Tris, pH 8.0, and 100 mM NaCl (Fig. 4). Protein binding to the column, washing, and on-column cleavage were carried out at 4 °C. Fractions were collected after an overnight incubation at 4 °C. Hpa I<sub>210</sub> protein concentrations were determined using the Bio-Rad Protein Assay (Hercules, California). RNase A concentrations were calculated using an extinction coefficient at 277.5 nM of 9.8 mM<sup>-1</sup> cm<sup>-1</sup> (Lyles & Gilbert, 1991).

#### Purification of MYB and MXB

Full-length precursor proteins consisting of maltose binding protein–*Xcc* VMA intein (N454A)–chitin binding domain (MYB) and maltose binding protein–*Mxe* CysA (N198A) intein–chitin binding domain (MXB) were purified after induction and sonication, as described for thioester-tagged proteins, by applying the sonicated sample to a 10 mL column of amylose resin (New England Bio-

labs). Unbound proteins were washed from the column with 10 column volumes of Buffer A (see purification of thioester-tagged proteins). Bound proteins were eluted with a buffer of 50 mM Tris, pH 8, containing 100 mM NaCl and 10 mM maltose. Fractions were collected and protein concentrations were determined using the Bio-Rad Protein Assay.

#### Peptide synthesis

Peptide synthesis was performed on an ABI model 433A peptide synthesizer utilizing *FactMoc* chemistry (Fields et al., 1991) at a 0.085 mmol scale. Pretreated HMP (*p*-hydroxymethyl-phenoxymethyl) polystyrene resins (Applied Biosystems, Foster City, California) functionalized at 0.5 mmol/g was used in conjunction with Fmoc/NMP chemistry utilizing HBTU amino acid activation (Douraglou et al., 1984; Knorr et al., 1989). Fmoc amino acids were purchased from Applied Biosystems. Synthesis proceeded with a single coupling during each cycle. Peptide cleavage from the resin and simultaneous removal of side chain protecting groups was facilitated by the addition of cleavage mixture (Perkin-Elmer) consisting of 0.75 g phenol, 0.25 mL 1,2-ethanedithiol, 0.5 mL deionized H<sub>2</sub>O, and 10 mL TFA. The resin was flushed with nitrogen and gently stirred at room temperature for 3 h. Following filtration and precipitation into cold (0 °C) methyl-*t*-butyl ether, the precipitate in the ether fraction was collected by centrifugation. The peptide precipitate was vacuum dried and analyzed by mass spectrometry using a Perceptive Biosystems (Framingham, Massachusetts) MALDI-TOF mass spectrometer. Final purification was by HPLC using a Waters HPLC system with a Lambda-Max Model 481 multiwavelength detector (set at 214 nm), 500 series pumps, and automated gradient controller with a Vydac semi-preparative C18 column. Elution of the peptide was with a 60 min linear gradient of 6–60% acetonitrile (v/v) in an aqueous solution of 0.1% TFA (v/v).

#### Protein cleavage and ligation reactions

Cleavage of MYB and MXB. The precursor protein (1 mg/mL) was incubated overnight at 4 °C with or without a third reagent (50 mM) in 150 mM Tris, pH 8, containing 100 mM NaCl. Ligation reactions with MYB and MXB. The precursor protein (1 mg/mL) was treated as described for cleavage except that a 30-amino acid peptide (1 mM final concentration, NH<sub>2</sub>-CAYKTTQANKHIVACEGNPYVPVHFDA5V-COOH) was also included in the reaction. Ligation reactions after purification of thioester-tagged proteins. Lyophilized peptides (New England Biolabs) were added (to 1 mM final concentration) directly to the thioester-tagged protein freshly isolated from the chitin column. The reaction was allowed to proceed overnight at 4 °C. In the case of RNase A, RNase A<sub>109</sub> was subjected to concentration (3–4 h concentration time) using a Centrprep 3 column (Amicon, Beverly, Massachusetts) before addition of the peptide.

#### HPLC purification of ligated RNase A

The ligated RNase A was separated from small amounts of impurities by HPLC. The HPLC apparatus consisted of a GP-250 Gradient Programmer, P-500 pumps and LKB Control Unit UV-1 (Pharmacia, Uppsala, Sweden). The sample was dialyzed into phosphate buffer (15 mM phosphate buffer, pH 6.6) containing 50 mM NaCl. This was applied to a mono S column (1 mL bed volume;



Pharmacia) equilibrated in the same buffer. Ligated RNase A was eluted with a 50–600 mM NaCl gradient in phosphate buffer. Both RNase A purchased from the Sigma Chemical Co. (St. Louis, Missouri) and ligated RNase A eluted at 210 mM NaCl under these conditions.

#### Determination of ligation efficiencies

Ligation efficiencies were based on the comparison of densitometry measurements taken by scanning the SDS-PAGE gels of ligation reactions. Staining intensity of protein bands from scanned images was determined using the program NIH Image 1.59/ppc with subtraction of the intensity of an equivalent area with no visible staining. Percent ligation efficiencies were calculated using the equation:

$$LE = LP * 100 / (LP + UP) \quad (1)$$

where *LE* is the ligation efficiency, *LP* is the staining intensity of ligated product, and *UP* is the staining intensity of unligated protein.

#### Renaturation of RNase A

Following the ligation reaction, RNase A was subjected to renaturation as described previously (Lyles & Gilbert, 1991). Ligated RNase A was initially denatured by overnight dialysis at 4 °C against 100 mM Tris, pH 8 containing 2 mM EDTA, 140 mM DTT, and 6 M guanidine HCl. This was followed by dialysis into a solution of 50 mM Tris buffer pH 8 containing 100 mM NaCl, 0.2 mM GSSG, 1.0 mM GSH, and 2 mM EDTA (overnight dialysis at 4 °C, followed by another 1 h in fresh solution). Refolding was facilitated by the addition of protein disulfide isomerase (1 U/330 µg of RNase A, Sigma Chemical Co.) and allowed to proceed for 7 h at room temperature.

#### Enzymatic assays for RNase A and HpaI

The catalytic activity of RNase A was monitored by its ability to hydrolyze cytidine 2',3'-cyclic monophosphate (C>P, Sigma Chemical Co.) as described previously (Hodges & Merrifield, 1975). A typical assay consisted of the addition of ligated or truncated RNase A (0.61 µM) to differing concentrations of C>P in 100 mM Tris buffer, pH 8 containing 100 mM NaCl. Hydrolysis of C>P was monitored by a change in absorbance at 296 nm using extinction coefficients of 0.19 mM<sup>-1</sup> cm<sup>-1</sup> and 0.38 mM<sup>-1</sup> cm<sup>-1</sup> for C>P and hydrolyzed C>P, respectively (Lyles & Gilbert, 1991). The activity of HpaI was determined by its ability to digest Lambda DNA (New England Biolabs). Serial dilutions of ligated or hydroxylamine treated truncated HpaI, with the appropriate peptide added to 1 mM, were incubated with 1 µg of Lambda DNA for 1 h at 37 °C in a buffer of 20 mM Tris-acetate, pH 7.9, 10 mM magnesium acetate, 50 mM potassium acetate, 1 mM dithiothreitol, and 170 µg/mL BSA (total volume 30 µL). Digestion reactions were visualized on 1% agarose gels permeated with ethidium bromide. One unit of HpaI was defined as the amount of enzyme necessary to digest 1 µg of Lambda DNA in 1 h at 37 °C.

#### Acknowledgments

We thank Henry Paulus, Christopher Noren, Francine Perler, Shaorong Chong, John Baswell, Chun Guan, Richard Whittaker, William Jack, Mau-

rice Southworth, Melissa Scott, and Tom Muir for valuable discussions and assistance; Francine Perler and Maurice Smithworth for the clone of the *Msc* CysA intein; Suzanne Kreske for purified HpaI; the organic synthesis and sequencing divisions of New England Biolabs for their assistance with special requests; and Donald Comb for support and encouragement.

#### References

- Chung TK, Jackson DY, Burnier JF, Wells JA. 1994. Subtiligase: A tool for semisynthesis of proteins. *Proc Natl Acad Sci USA* 91:12544–12548.
- Chong S, Marsh FB, Comb DG, Scott ME, Landry D, Vance LM, Perler FB, Banner T, Kucera RB, Hirvonen CA, Pellerin JJ, Paulus H, Xu M-Q. 1997. Single-column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element. *Gene* 192:271–281.
- Chong S, Shao V, Paulus H, Besser J, Perler FB, Xu M-Q. 1996. Protein splicing involving the *Saccharomyces cerevisiae* YMA intein. *J Biol Chem* 271:22159–22168.
- Chong S, Williams KS, Wolkowicz C, Xu M-Q. 1998. Modulation of protein splicing of the *Saccharomyces cerevisiae* vacuolar ATPase (VMA) intein. *J Biol Chem* 273:10567–10577.
- Dawson PE, Moir TW, Clark-Lewis I, Kent SB. 1994. Synthesis of proteins by native chemical ligation. *Science* 266:776–779.
- Doumleau V, Gross B, Lanthropoulos V, Zomdrou C. 1984. O-benzotriazoly-N,N,N',N'-tetraacetyluroniumhexafluorophosphate as coupling reagent for the synthesis of peptides of biological interest. *Synthesis* 372–374.
- Duplay P, Badouelle H, Fowler A, Zabin J, Saurin W, Hofnung M. 1984. Sequences of the malE gene and of its product, the malE-binding protein of *Escherichia coli* K12. *J Biol Chem* 259:10606–10613.
- Eklund DA, Chytil M, Verdine GL. 1996. The leucine zipper domain controls the orientation of AP-1 in the NF-κB/AP-1-DNA complex. *Chem Biol* 3:981–991.
- Felds CG, Lloyd DM, MacDonald RL, Oteson KM, Noble RL. 1991. HBTU activation for automated Fmoc solid-phase peptide synthesis. *Pept Res* 4:95–101.
- Guan C, Li F, Biggs PD, Inouye H. 1987. Vectors that facilitate the expression and purification of foreign peptides in *Escherichia coli* by fusion to maltose-binding protein. *Gene* 67:21–30.
- Guthe B, Lin M, Calhoun D, Merrifield R. 1972. Reactivation of des(119–, 120– or 121–124) ribonuclease A by mixture with synthetic COOH-terminal peptides of varying lengths. *J Biol Chem* 247:4763–4767.
- Hodges RS, Merrifield R. 1975. The role of serine-123 in the activity and specificity of ribonuclease. *J Biol Chem* 250:1231–1241.
- Ito H, Shimoto H, Sadakata A, Kawanishi H, Kimizuka F, Kato I. 1992. Cloning and expression of the HpaI restriction-modification genes. *Nucleic Acids Res* 20:705–709.
- Jackson DY, Burnier J, Quas C, Stanley M, Iben J, Wells JA. 1994. A designed peptide ligase for total synthesis of ribonuclease A with unnatural catalytic residues. *Science* 266:243–247.
- Kane PM, Yamashiro CI, Whiczyk DE, Neff N, Goehi M, Stevens TH. 1990. Protein splicing converts the yeast TFP1 gene product to the 69-kD subunit of the vacuolar H<sup>+</sup>-adenosine triphosphatase. *Science* 250:651–657.
- Kent SB. 1988. Chemical synthesis of peptides and proteins. *Annu Rev Biochem* 57:957–989.
- Knor R, Trzeciak A, Bannwarth W, Gillieson D. 1989. New coupling reagents in peptide chemistry. *Tetrahedron Lett* 30:1937–1939.
- Lin M. 1970. The structural roles of amino acid residues near the carboxyl terminus of bovine pancreatic ribonuclease A. *J Biol Chem* 245:6726–6731.
- Liu CF, Tam JP. 1994. Peptide segment ligation strategy without use of protecting groups. *Proc Natl Acad Sci USA* 91:6584–6588.
- Lu W, Qasim MA, Kent SBH. 1996. Comparative total syntheses of turkey ovomucoid third domain by both stepwise solid phase peptide synthesis and native chemical ligation. *J Am Chem Soc* 118:8518–8523.
- Lyles MM, Gilbert W. 1991. Catalysis of the oxidative folding of ribonuclease A by protein disulfide isomerase: Dependence of the rate on the composition of the redox buffer. *Biochemistry* 30:613–619.
- Nakagawa SH, Tager HS. 1986. Role of the phenylalanine B25 side chain in directing insulin interaction with its receptor. *J Biol Chem* 261:7332–7341.
- Noren C, Anthony-Cahill S, Griffith M, Schultz E. 1989. A general method for site-specific incorporation of unnatural amino acids into proteins. *Science* 244:182–188.
- Offord R. 1987. Protein engineering by chemical means? *Protein Eng* 1:151–157.
- Perler FB, Davis EO, Dean GE, Gimble FS, Jack WE, Neff N, Noren CJ, Thomsen J, Belfort M. 1994. Protein splicing elements: Inteins and exteins—A definition of terms and recommended nomenclature. *Nucleic Acids Res* 22:1125–1127.

- Roy RH, Acharya AS. 1993. Semisynthesis of hemoglobin. *Methods Enzymol* 211:194-215.
- Saiki RK, Schaff S, Faloutsos F, Mullis KB, Horn GT, Erlich HA, Arnheim N. 1985. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* 230:1350-1354.
- Shao Y, Xu M-Q, Paulus H. 1997. Protein splicing: Characterization of the aminomucinimide residue at the carboxyl terminus of the excised intervening sequence. *Biochemistry* 36:10847-10850.
- Tillett A, Southworth M, Akende P, Douglas S, Jacobs I, William R, Perler FB. 1997. The *Mycobacterium xenopi* QyrA protein splicing element: Characterization of a minimal motif. *J Bacteriol* 179:6378-6382.
- Wallace CJ. 1993. Understanding cytochrome c function: Engineering protein structure by semisynthesis. *FASEB J* 7:505-515.
- Wallace CJ. 1995. Peptide ligation and semisynthesis. *Curr Opin Biotechnol* 6:403-410.
- Watanabe T, Ito Y, Yamada T, Hashimoto M, Sekine S, Tanaka H. 1994. The role of the C-terminal domain and type III domain of chitinase A1 from *Bacillus circulans* WL-12 in *in vitro* degradation. *J Bacteriol* 176:4465-4472.
- Wlodawer A, Svensson L, Sjölin L, Gilliland G. 1988. Structure of phosphatidylesterase A refined at 1.26 Å. *Biochemistry* 27:7703-7717.
- Woods AC, Gollan-Smith KL, Parish JC, Smith M, Wallace CJ. 1996. Synergy in protein engineering: Mutagenic manipulation of protein structure to simplify semisynthesis. *J Biol Chem* 271:32008-32015.
- Xu M-Q, Comb DG, Paulus H, Morris CJ, Shao Y, Perler FB. 1994. Protein splicing: An analysis of the branched intermediate and its resolution by succinimide formation. *EMBO J* 13:5517-5522.
- Xu M-Q, Perler FB. 1996. The mechanism of protein splicing and its modulation by mutation. *EMBO J* 15:5146-5153.

New England Biolabs, Inc.  
2003-2004 Catalog  
Pages 164-166

## IMPACT<sup>®</sup>-CN System

recombinant fusion protein expression and one-step affinity purification

#E6600S ..... \$400

### Components Sold Separately:

#### Chitin Beads

#S6651S 20 ml ... \$45

#S6651L 100 ml ... \$180

#### Anti-Chitin Binding Domain Serum

#G6604S 0.05 ml ... \$50

### Vectors Sold Separately:

See page 166 for a complete listing and guide to IMPACT vectors. Restriction maps can be found on pages 316-317; sequence files are available at <www.neb.com>.

### Primers Sold Separately:

See page 140 for a complete listing of IMPACT sequencing primers.

### The System Includes:

- Four pTYB E. coli expression vectors: pTYB1, pTYB2, pTYB11, pTYB12
- Control Plasmid (pMYD5)
- Chitin Beads: Affinity matrix used for fusion protein purification (binding capacity = 2 mg/ml)
- E. coli Strain ER2686
- Anti-CBD Serum (rabbit)
- T7 Universal, T7 Terminator Reverse, Intron Forward and Intron Reverse Sequencing Primers
- DTT and SDS Sample Buffer
- Comprehensive Instruction Manual

**Description:** IMPACT<sup>®</sup> (Isletin Mediated Purification with an Affinity Chitin-binding Tag) utilizes the inducible self-cleavage activity of engineered protein splicing elements (termed isletins) to purify recombinant proteins by a single affinity column (1-4) (Figures 1 and 2). This system distinguishes itself from other protein fusion systems by its ability to separate a recombinant protein from the affinity tag without the use of a protease.

The IMPACT-CN Protein Fusion and Purification System contains four expression vectors (pTYB vectors) which allow fusion of a bifunctional tag, consisting of the isletin and the chitin binding domain (5), to either the C-terminus (pTYB1,2) or N-terminus (pTYB11,12) of the target protein. In the presence of isletins such as DTT,  $\beta$ -mercaptoethanol or cysteine, the isletin undergoes specific self-cleavage which releases the target protein from the chitin-bound isletin tag.

Both pTYB1 and pTYB11 contain a Sap I cloning site which allows the target gene to be cloned immediately adjacent to the cleavage site of the isletin tag; this results in the purification of a target protein without any extra non-active residues attached to its terminus after cleavage. Use of pTYB2 or pTYB12 yields a target protein with extra residue(s) added in its C-terminus or N-terminus, respectively. Addition of extra residues may help the cleavage reaction for some target

proteins. pTYB2 and pTYB12 contain the same compatible restriction sites in the multiple cloning region. This allows cloning the same target gene fragment into both vectors. pTYB1 and pTYB2 use the Nde I site in the multiple cloning region for translation initiation.

Vectors sold separately: pTYB3 and pTYB4 contain a Neo<sup>r</sup> site in place of the Nde I site in pTYB1 and pTYB2, respectively. pTYB5 is identical to pTYB1 except that it carries kanamycin resistance instead of ampicillin resistance. pTYB6 and pTYB7 contain mini-isletin with alternative preferred residues at cleavage site (10,11). See page 166 for a guide to IMPACT vectors and applications.

### References:

- (1) Chung, S. et al. (1997) *Gene* 192, 277-287
- (2) Chung, S. et al. (1998) *Nucl. Acids Res.* 26, 5110-5115.
- (3) Chung, S. et al. (1998) *J. Biol. Chem.* 273, 10667-10677.
- (4) Chung, S. et al. (1996) *J. Biol. Chem.* 271, 22153-22158.
- (5) Watanabe, T. et al. (1994) *J. Bacteriol.* 176, 4472.
- (6) Duhendorf, J. W. and Studier, F. W. (1990) *Mol. Biol.* 219, 45-59.

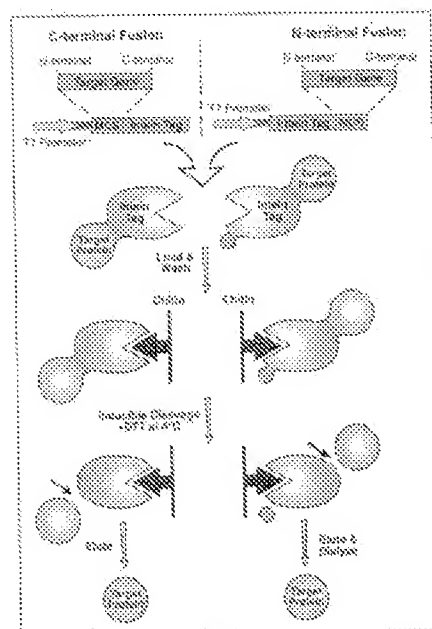


Figure 1: Schematic of the IMPACT-CN System.

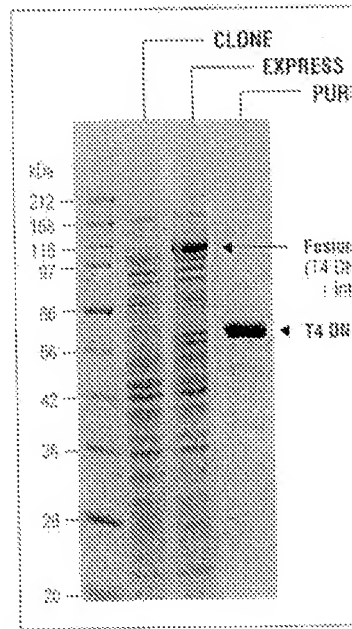


Figure 2: Purification of biologically active T4 Ligase with >88% purity in one step. Sd: NE. Range Protein Marker. Lane 1 (Clone): undisturbed extract. Lane 2 (Express): induced cell extract expressing fusion protein. Lane 3 (Purify): T4 fractions eluted after inducing cleavage over-

## IMPACT™-TWIN System

protein labeling, ligation and cyclization

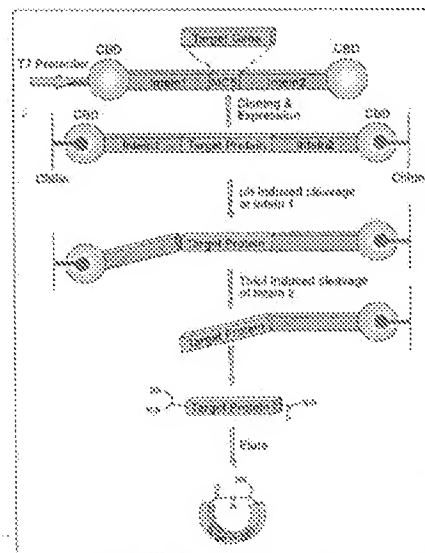
**Description:** The IMPACT™-TWIN Protein Fusion and Purification System utilizes the inducible self-cleavage activity of engineered protein splicing elements (termed inteins) for protein purification and manipulation. It can produce a target protein with an N-terminal cysteine and/or a C-terminal thioester for protein labeling, ligation and cyclization.

The System contains three expression vectors that allow the fusion of a target protein to modified intein tags. These vectors permit the isolation of proteins with a C-terminal thioester for protein ligation or with an N-terminal residue other than Met (7-9). The protein ligation reaction allows the incorporation of non-coded amino acids or labels into a protein sequence. Furthermore, pTWIN vectors can be used to fuse a target protein to the C-terminus of intein1, which can be induced to undergo cleavage by a temperature/pH shift (5). This is advantageous for purifying target proteins that are sensitive to reducing agents. Finally, insertion of the target protein between two intein tags allows the generation of circular protein species with a peptide bond at the site of ligation (7).

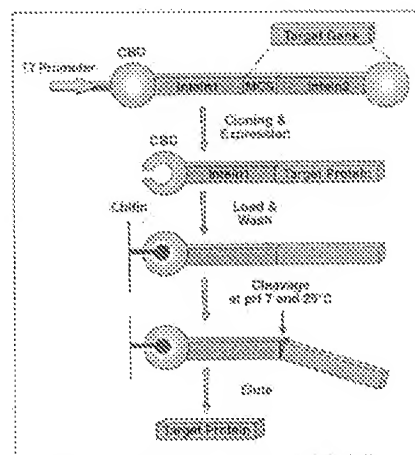
**Note:** The controlled cleavage activity of an intein tag can be affected by the target protein sequence. Please see Table 1 (page 186) for amino acids that are most likely to cleave as expected. Also, when creating circular proteins it is typical that the linear form will be co-purified, requiring a further step to separate the two protein species.

## References:

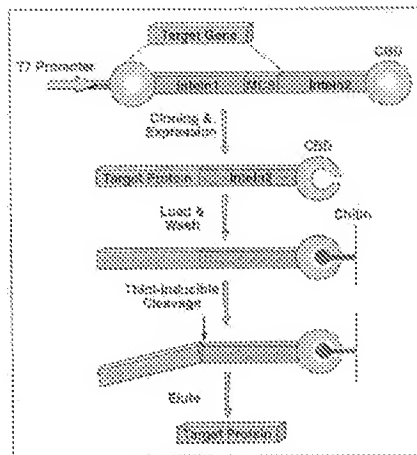
- (7) Evans, T. C. et al. (1999) *J. Biol. Chem.* 274, 18353-18363.
- (8) Evans, T. C. et al. (1999) *J. Biol. Chem.* 274, 3923-3926.
- (9) Mathys, C. et al. (1995) *Gene* 231, 1-13.
- (10) Evans, T. C. et al. (1998) *Protein Sci.* 7, 2256-2264.
- (11) Southworth, M. et al. (1998) *Biochemistry* 27, 110-120.



**Figure 1: Protein cyclization:** The target protein is fused between Intein1 and Intein2 and released with an N-terminal cysteine and a C-terminal thioester.



**Figure 2: Protein purification by pH and temperature induced cleavage:** Using a pTWIN vector, the intein tag is fused to the N-terminus of the target protein. The target protein is then released by pH and temperature induced cleavage.



**Figure 3: Protein purification by thiol-induced cleavage:** Using a pTWIN vector, the intein tag is fused to the C-terminus of the target protein. The target protein is then released by thiol-induced cleavage.

#EE9508 ..... \$400

## Components Sold Separately:

## Chitin Beads

#S6651S 20 ml ... \$45  
#S6651L 100 ml ... \$180

## Anti-Chitin Binding Domain Serum

#S66540 0.05 ml ... \$50

## Vectors Sold Separately:

See page 166 for a complete listing and guide to IMPACT vectors. Restriction maps can be found on pages 310-317; sequence files are available at <www.neb.com>.

## Primers Sold Separately:

See page 140 for a complete listing of IMPACT sequencing primers.

## The System Includes:

- Three pTWIN E. coli expression vectors: pTWIN1, pTWIN2, pTWIN-MBP1
- Chitin Beads: Affinity matrix used for fusion protein purification (binding capacity=2 mg/ml).
- E. coli strain ER2536
- Anti-CSD Serum (rabbit)
- Ssp DnaB Intein Forward, Mxe Intein Reverse and Mtn hsp1 Intein Reverse Sequencing Primers
- DTT and SDS-PAGE Sample Buffer
- Instruction Manual

## Advantages of the IMPACT Systems:

- Single-column purification—no additional steps to remove the affinity tag
- Flexibility of fusion to either the C-terminus or N-terminus of the target protein maximizes the probability of successful expression and purification
- Yields native amino acid sequence
- Release of fusion partner without the use of proteases
- Isolation of proteins with or without an N-terminal methionine residue
- Use of either thiol reagent or pH and temperature shift to induce on-column cleavage
- The only system that produces proteins possessing an N-terminal cysteine and/or C-terminal thioester for use in protein labeling, ligation and cyclization
- IPTG-inducible T7 Promoter (5) for high-level expression and tight regulation of expression



## IMPACT™ Vectors

Vectors	Insert length (amino acids)	Target protein fusion site	Preferred residues at cleavage site	Cleavage method	Applications
<b>IMPACT-CN KH</b> pTYB1.2	454	C-terminus	All except P C N D R	Thiol	Purification (1); C-terminal thioester for gel
pTYB1.1,2	510	N-terminus	All except P S G T	Thiol	Purification (2)
<b>IMPACT-TWIN KH</b> pTWIN1	154 198	N-terminus C-terminus	S, C, G, A, M, I All except S P E D G A	pH & temperature Thiol	Purification; C-terminal thioester, defined to e.g. Cys; protein ligation and cyclization (1)
pTWIN2	154 198	N-terminus C-terminus	S, C, G, A, M, T S, A	pH & temperature Thiol	Purification; C-terminal thioester, defined to e.g. Cys; protein ligation and cyclization (1)
<b>Sold Separately</b> pTYB2.4 and pTYB1	454	C-terminus	All except P C N D R	Thiol	Purification; C-terminal thioester for protein
pTXB1.3	198	C-terminus	All except S P E D G A	Thiol	Purification (10,11); C-terminal thioester for gel

Table 1: Guide to IMPACT vectors and applications.

## IMPACT Vectors Sold Separately:

<b>pTYB1 Vector DNA</b> #N6701S	10 µg ....	\$75
<b>pTYB2 Vector DNA</b> #N6702S	10 µg ....	\$75
<b>pTYB11 Vector DNA</b> #N6901S	10 µg ....	\$75
<b>pTYB12 Vector DNA</b> #N6902S	10 µg ....	\$75
<b>pMYB5 Control Plasmid</b> #N6906S	10 µg ....	\$50
<b>pKYB1 Vector DNA</b> #N6706S	10 µg ....	\$75
<b>pTXB1 Vector DNA</b> #N6707S	10 µg ....	\$75
<b>pTXB3 Vector DNA</b> #N6708S	10 µg ....	\$75
<b>pTYB3 Vector DNA</b> #N6703S	10 µg ....	\$75
<b>pTYB4 Vector DNA</b> #N6704S	10 µg ....	\$75
<b>pTWIN1 Vector DNA</b> #N6911S	10 µg ....	\$75
<b>pTWIN2 Vector DNA</b> #N6912S	10 µg ....	\$75
<b>pTWIN-MBP1 Vector DNA</b> #N6955S	10 µg ....	\$75

Note: *E. coli* strain ER2566 is included with all vectors sold separately.

Notice to Buyer/User: Non-profit buyers or users have a non-exclusive license to use this system or any component thereof for RESEARCH PURPOSES ONLY. Non-profit buyers or users must have or obtain advance for RESEARCH PURPOSES from Brockton National Laboratory, Building 470D, P.O. Box 1000, Upton, MA 01569-9000; telephone: 508-534-7134; fax: 508-534-3726. For commercial use of this system or any components thereof, with Non-profit and For-profit buyers and users may obtain a license from New England Biolabs. Additional information regarding licensing of this system may be obtained from NEB.

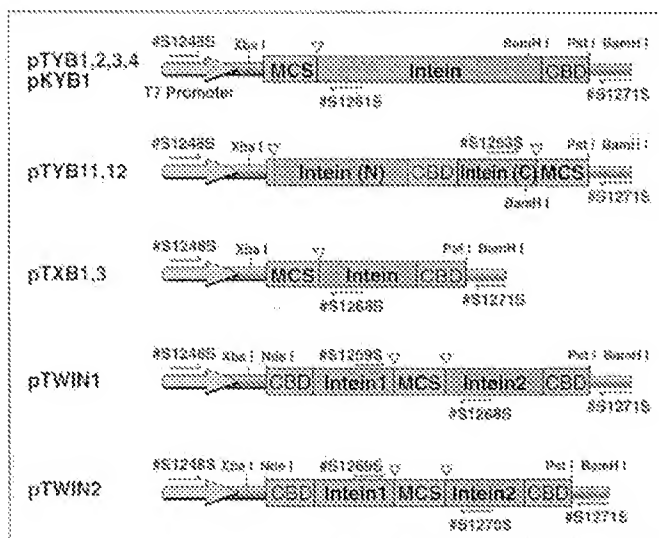


Figure 1: IMPACT vectors and sequencing primers. Detailed restriction maps can be found on pages 310-317.

CBD = chitin binding domain, MCS = multiple cloning site, ▽ = intein cleavage site(s)